DOCUMENT RESUME

ED 198 570 CS 503 135

AUTHOR Hensley, Wayne E.: Sanford, David L.
TITLE Personality as a Determinate of Response Dimension Scaling for Likert Rating Categories.
PUB DATE Apr 81
NOTE 26p.: Paper presented at the Annual Meeting of the Southern Speech Communication Association (Austin, TX, April 8-10, 1981).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Anxiety: Attitude Measures: Behavioral Science Research: College Students: Communication Problems: Communication Research: Higher Education: *Personality Traits: Rating Scales: *Research Problems: *Response Style (Tests): *Scaling: Self Esteem: *Semantic Differential: Social Science Research
IDENTIFIERS Communication Apprehension: *Likert Scales

ABSTRACT

A study was designed to evaluate the use of summated rating (Likert) scales of agreement, evaluation, and frequency. The subjects, 58 female and 45 male college students, rank ordered the descriptive adjectives for the areas of agreement, evaluation, and frequency on a scale of 1 to 100. They also completed the Personal Report of Communication Apprehension (PRCA) and the Rosenberg self-esteem measure. The analysis of the subjects' ratings of the three sets of descriptive adjectives supported earlier research by P. E. Spector, indicating that the assumptions of equal interval scaling were untenable. Thus, the differences between the "a priori" method of selecting standard scale response values and the actual scales used by each subject produce important differences in the interpretation of social science research. Another finding was that different personality types used the response categories of agreement, evaluation, and frequency in significantly different ways. Subjects with low self-esteem and high apprehension were in less agreement about the meaning of the terms than the subjects with high self-esteem and confidence. The findings also suggested that the center regions of the scales were more stable than the ends, and that the negative ends of the scales might be more unstable than the positive ends. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED198570

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Personality as a Determinate of Response Dimension Scaling for Likert Rating Categories

Wayne E. Hensley

and

David L. Sanford

Virginia Polytechnic Institute

and State University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Wayne E. Hensley

David L. Sanford

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Competitively selected paper presented to the Interpersonal Communication interest group at the Southern Speech Communication Association annual convention, Austin, Texas, April 8-10, 1981.

C SS83/35

Personality as a Determinate of Response Dimension Scaling for Likert Rating Categories

A great deal of knowledge in the area of social science research rests on the foundation of summated rating (Likert) scales. We ask persons whether they strongly agree or disagree with a public policy; whether they rarely or frequently perform a certain activity; whether they think a proposal is terrible or excellent. Based on the answers to such queries, responses are measured and inferences are made concerning attitudes and behavior.

While it is not universally the case, the results of summated rating scales are almost invaribly treated as interval data. That is, the basic assumption has been that the distance between "Strongly agree" and "Agree" has been presumed to be the same as the distance between "Undecided" and "Disagree." This assumption allows the addition and subtraction of scale values to form a composite index. In addition, the interval assumption allows the application of parametric statistics with rather sophisticated implications being possible. Unfortunately, the equal interval assumption is often made on the basis of no better foundation than the personal experience of the researcher or simply convenience. A recent study (Spector, 1976) raises the strong possibility that the widely held assumption of equal intervals may not be accurate.

Spector (1976) provided 107 college undergraduates with a list of descriptive adjectives for the response categories of agreement, evaluation, and frequency. The respondents were asked to rank order each of the three lists; no ties were allowed. The obtained results were disquieting. For example, ranked in a field of 13 adjectives, "Halfway agree" was not halfway; it was actually 4.8. "Inclined to agree" and "Moderately agree" were given identical ranks (5.4) with both differing substantially from "Pretty much agree" (7.0). Similar examples can be drawn from the other scales of evaluation and frequency. In short, the entire equal interval assumption is open to question.

There are at least two distinct possibilities to explain Spector's results. First, he may be correct; the equal interval assumption may have been invalid for all these years. On the other hand, the ranking method per se may be the culprit by forcing a spurious order among isomorphic concepts. Consider for example, A, B, and C to be essentially identical descriptions among a field of 13 total terms. If there is truly no way for respondents to distinguish among A, B, and C, we would expect them to be randomly ranked within a single judgmental bracket. In addition however, we would also expect the interquartile range of the ranks to overlap substantially among the three. The only way to convince a skeptic that A, B, and C are distinct would be to demonstrate the consistency of the rank order among the terms. Spector did not do this.

A more convincing test of the equal interval assumption applied to summated rating scales is to give respondents a true interval scale and ask them to place "Halfway agree," "Inclined to agree," and "Moderately agree" on the scale itself. If it is the case that certain adjectives are equivalent, then the equivalency

may be demonstrated. If there are equal intervals among the terms, such intervals may be directly observed. But the equal interval assumption can not be tested using an ordinal scale.

A second problem with the Spector study was that he used only half of the agreement scale, i. e., only the word "Agree" and phrases that modify that word, such as "Moderately agree" and "Strongly agree," were used. This, despite the standard practice of representing the agreement scale as ranging from some anchor endpoint like "Strongly agree" to the other anchor endpoint of "Strongly disagree." Although it could be argued that the "Disagree" part of the scale would simply be a mirror image of the "Agree" part, without testing this is pure conjecture. In addition, Spector does not settle the issue concerning the interpretation of midrange terms such as "Undecided" and "Indifferent" since he left them out entirely.

What seems to be needed is an omnibus study of the response categories of agreement, evaluation, and frequency. The interval scale provided for judgments must allow the full range of responses to be assessed for each response category. This undertaking is the initial thrust of the present study.

Operating from another vantage point, it is well documented that different people see the world in very different ways depending on such variables as environment (Hastorf & Cantril, 1954), experience (Ittelson & Kilpatrick, 1951), and learning (Allport & Pettigrew, 1957). By extension, different orientations might also be expected from individuals who differ markedly on personality variables since personality is, in a sense, the "lens" through which reality is perceived. Thus, we might reasonably expect those with differing personality profiles to see response categories from differing perspectives.

Certainly one of the most central concepts of personality is self-esteem. It would be difficult to imagine an assessment of personality which did not include some mention of an individual's self-appraisal. To underscore the importance of the concept, we need only observe that there have been over two thousand publications pertaining to the self (Gordon & Gergen, 1968), and over two hundred scales to measure self-esteem (Robinson & Shaver, 1969). Clearly, it is a variable worth considering in any examination of personality.

Of the many measures used to assess self-esteem, one seems uniquely suited to most research using college students. The Rosenberg scale of self-esteem was developed specifically for use among young adults (Rosenberg, 1965). This measure has a definable unidimensional structure among college aged persons (Hensley & Roberts, 1976) and has been shown to be stable across different samples of college students (Hensley, 1977). In addition, the scale has typically been administered using the common response categories: "Strongly agree," "Agree," "Undecided," "Disagree," and "Strongly disagree."

Perhaps an equally important personality orientation--at least to researchers in communication or social psychology--are the communication attributes of the respondents themselves. If we observe that oral communication is by far the most common form of everyday human interaction, then any personality orientation tending to disrupt that interaction is of obvious interest. Beginning with the original report in the early 1970's (McCroskey, 1970), a very large amount of research has been conducted defining communication apprehension as a major

personality variable (McCroskey, 1977). In addition, the reliability and validity of this unidimensional measure are now well documented (McCroskey, 1978). Finally, this Personal Report of Communication Apprehension (PRCA) also uses the five response categories listed above.

Having observed the potential importance of personality in terms of individual responses, the final problem with the Spector study is now apparent. Although he was concerned about the equal interval assumption and the potential effects of violating this assumption in current social science research, Spector did not administer any standard rating scales to his respondents. Therefore, he could not evaluate whether the unequal intervals he found were a sufficient deviation from the hypothesized ideal that they affected the interpretation of social science data. In short, do the differences in the Spector study make any real differences?

What is proposed then is a replication and extension. First the original study of response category choices (Spector, 1976) will be refined and replicated using an interval rather than an ordinal scaling procedure. Second, the response category choices will be examined in light of the two personality variables of self-esteem and communication apprehension.

Method

The subjects were 58 female and 45 male undergraduate volunteers from several Communication courses at a major southeastern American university. The average age among respondents was 20.25 years, ranging from 18 to 32 years of age with a standard deviation of 2.14 years. The average grade point was 2.55 on a 4 point scale, ranging from 1.08 to 3.94 with a standard deviation of 0.58.

A two-part questionnaire was administered to all subjects. Following the introduction of the study and collection of the demographic data summarized above, the first portion of the questionnaire presented the descriptive adjectives for the areas of agreement, evaluation, and frequency with each area appearing on a separate page. For each judgmental area, a visual scale was presented for subjects to use as a reference:

1_____100

Depending on the area being represented, the anchor ends (1 and 100) of the visual reference were labeled for the subjects. The labels were: "ABSOLUTELY NEVER" and "ABSOLUTELY ALWAYS" (frequency), "STRONGEST POSSIBLE AGREEMENT" and "STRONGEST POSSIBLE DISAGREEMENT" (agreement), and "WORST POSSIBLE EVALUATION" and "BEST POSSIBLE EVALUATION" (evaluation). Below this visual image of each reference scale was the list of response terms that applied to that scale area. Each page asked the subject to put a number ranging from 1 to 100 beside each of the descriptive adjectives to indicate where on the scale that term fell.

After completing and turning in the first part of the questionnaire, the subjects were given the second part consisting of the 25 item PRCA and the 10 item Rosenberg self-esteem measure. Both the PRCA and the self-esteem measures were answered by coding the responses directly on to an optical scan computer form. All subjects completed both parts of the questionnaire.

Results and Discussion

The initial concern in this study was to ascertain whether or not equal intervals are perceived by respondents using the standard descriptors for scales of agreement, evaluation, and frequency. Since each of these dimensions constitutes a separable issue, we will consider them one at a time in the next three parts. The last two parts of the Results and Discussion will extend the analysis further.

Agreement scale values

The correspondence between our results and Spector's may be compared by observing the mean values awarded to agreement descriptors as shown in the overall column of Table 1. Since Spector had his terms rank ordered, his scale values ranged from 1 to 13 while our scale values range from 1.0 to 100.0. Given the different techniques used, some of the results are surprisingly similar. Spector found that "Very much agree" was given a rank of 9.1, "Pretty much agree" was ranked 7.0 while "Generally agree" scaled at 6.8.

Insert Table 1 about here.

The corresponding scale values for the same items, shown in column 1 of Table 1, are 88.71, 75.94 and 71.09. This is a reasonably close correspondence. Consider, however, the terms "Moderately agree," "Inclined to agree," "Tend to agree," and "Agree in part." Spector scores them at 5.4, 5.4, 5.3, and 4.7 respectively. This would suggest that although the first three terms are basically the same, the fourth is seen differently, being almost three quarters of a rank away from the others. Our interval scores for those same terms are 64.93, 66.51, 65.13 and 65.69 respectively. The use of an interval scale illustrates the indistinguishability of the terms since they differ by less than two points of a 100 point scale. In perspective however, after granting the greater precision available with an interval scale, the results of the present study and Spector's on the agreement dimension are monotonically quite similar ($r_s = .94$; $N = 13$; $p < .001$).

But it was the equal interval assumption which interested Spector the most, and it also is the main concern here. The spacing between common category terms can be examined by inspection of Table 1. The set most commonly used in agreement scaling provides an interesting example; "Strongly agree" is 90.95, "Agree" is 83.95, "Undecided" is 49.44, "Disagree" is 18.10, and "Strongly disagree" is 9.49. Thus, while the interval between "Strongly agree" and "Agree" is 7.00, the interval between "Undecided" and "Disagree" is 31.34. On the basis of these data, it cannot be argued that the equal interval assumption is supported.

Evaluation scale values

Again there is good general agreement on the descriptors using Spector's ranks and the interval scale values of the present study ($r_s = .97$; $N = 13$; $p < .001$) with the same sort of exceptions noted earlier. The ranks for "Decent" and "Fair"

were 6.0 and 6.1 for example. On an interval scale "Decent" was 63.10 while "Fair" was 56.44.

Once more, however, it is the interval assumption which is of prime concern. At our university, the words used to describe letter grades are: "Excellent," "Good," "Fair," "Barely passing," and "Failure." The scale values associated with these words are 94.91, 74.80, 56.44, 53.24, and 27.10.[1] The interval assumption is once more called into question.

Insert Table 2 about here.

Frequency scale values

The frequency descriptors have the strongest correspondence between the two studies (r_s = .99; N = 13; p < .001). Depending on the descriptors chosen for scale values, it would be possible to construct an almost equal interval scale for the frequency dimension. The irony is that most social scientists do not typically use frequency scales. The relative purity of these judgments is therefore interesting but not terribly useful.

Insert Table 3 about here.

In summary, the equal interval assumption cannot be supported in the case of agreement scaling; it is open to serious question in evaluation scaling; finally, it is problematic in frequency scaling. This portion of the analysis reveals a corroboration of the earlier research (Spector, 1976) and offers a further refinement of that study. Based on these data, we conlcude that, for the instances of greatest interest to the social scientist, the assumptions of equal interval scaling are untenable.

Beyond Spector--One

The next step in the analysis begins to investigate the consequences of using scale response terms of unequal intervals. There were two ways this could have been done. The first would be to simply recode all subjects' "Strongly agree" responses from 5 to 90.95, all "Agree" responses from 4 to 83.95, etc., and to compare the typical 5-4-3-2-1 distribution of responses with the recoded distribution of mean responses. However, this would only examine the differences between a measure standardized by the dogma of infallible authority and a measure standardized by the dogma of refined knowledge.

The second alternative, and the one we chose, was to take each subject's

answers to the PRCA and self-esteem sections of the questionnaire and recode them using that subject's score for each response term taken from their answers to the agreement scale section of the questionnaire. This technique has the slight disadvantage of obscuring the issue about unequal interval widths. After all, some subjects assigned smaller values to "Disagree" than other subjects assigned to "Strongly disagree." This results in having some subjects' "Disagree" interpreted as more disagreement than other subjects' "Strongly disagree." Moreover, many subjects gave, for example, "Disagree" a score of 1 and "Strongly disagree" a score of 10, as well as giving "Agree" a score of 100 and "Strongly agree" a score of 90. Therefore, we are not only changing interval widths, but in many cases we are changing the order of responses. It is hoped that the reader agrees with the authors that this disadvantage is worth the attempt to broaden the perspective to include all available information on each subject's unique cognitive structuring of the research task. We are still asking if the differences make a difference, only now we are including more complex differences.

This scheme yielded four sets of data: the usual PRCA measure data, the recoded PRCA data, the usual self-esteem measure data, and the recoded self-esteem data. There are a number of ways to examine the differences and similarities among sets of data. One of the major choices among alternative techniques is the question of using descriptive or predictive statistical techniques. All of part One concentrates on what can be considered descriptive techniques of examining the four data sets, while part Two concentrates on predictive methods.

One of the most common ways of describing the underlying structure of a data set is by factor analysis. Factor analysis includes a large variety of techniques, but generally involves the factoring of the data to produce a smaller data matrix which accounts for or describes all the significant variability in the original data, and then rotating this factor solution into some kind of final form, either oblique or orthogonal. The most commonly used set of techniques currently chosen is Kaiser's original "Little Jiffy" solution: principal components factoring followed by varimax rotation. Recent studies (Acito & Anderson, 1980) indicate that this solution is not appropriate for small samples, i. e., around 100 subjects. Realizing that factoring procedures generally produce extra factors when the sample size is as small as in this study, we chose to use alpha factoring over image factoring, since image factoring always gives extra factors anyway. And since, as mentioned previously, both the PRCA and Rosenberg measures are so familiar and assumed to possess stable, unidimensional structures, we decided to stay with varimax rotation. The result was that both the usual self-esteem data and the recoded self-esteem data produced 2-factor solutions, as can be seen in Table 4. However, the usual PRCA data produced a 5-factor solution

---

Insert Table 4 about here.

Insert Table 5 about here.

Insert Table 6 about here.

(Table 5), while the recoded PRCA data produced a 7-factor solution (Table 6). At least with the PRCA measure, there are observable differences in the data sets. While we are clearly not attempting any definitive description of the PRCA in this paper, these results tend to corroborate recent studies of that measure. In two separate investigations (Beatty, Behnke, & McCallum, 1978; Parks, 1980) the PRCA has been found to be a situationally unstable scale. It may be speculated that the greater refinement of interval measurement used here is picking up some of this inherent instability.

The fact that they are different does not come as any surprise, since we were aware of some of the differences before the factor analyses were computed. We sought another approach to describing the differences among the four data sets. In a sense, we have been testing the ability of standardized response values on instruments such as the PRCA and Rosenberg measures to categorize unstandardized human thought processes. Therefore we thought that an important question to ask is which version of the data sets minimize the measurement error of the instruments, i. e., which were more reliable measures of communication apprehension and self-esteem. Cronbach's standardized item alpha reliability coefficient was computed for all four data sets. One should recall that Cronbach's alpha has a maximum value of 1.0 and that the closer the computed coefficient comes to 1.0, the less measurement error the data set has. The alpha value for the usual PRCA data is 0.099, while the recoded PRCA data has a value of 0.680. Similarly, the usual self-esteem data's alpha value is -0.765, while the recoded self-esteem has a value of 0.439. In both sets of data sets, the recoded data has considerably less measurement error than their typical standardized counterparts, that is, using each subject's personal scale values produces a considerably better measure of their apprehension and esteem than using the standard scale values.

Beyond Spector--Two

The differences in factor structures and reliability coefficients discussed in the previous section tend to support the hypothesis that the differences between the a priori method of selecting standard scale response values and the actual scales used by each subject produce important differences in the interpretation of social science research. Unfortunately, in the growing wealth of differences it is easy to lose sight of which differences to focus on as the most important and interesting. In this section, we identify and investigate the differences we concluded were the most important, namely, the differences among the subjects' scales.

Which measure(s)?: It would be possible to examine subject reactions in two different modes. We have two versions of the PRCA and of the Rosenberg

self-esteem measures--the usual version and the recoded version. However, since the concern here is in examining and comparing the subjects' response scales, one version each of the self-esteem measure and the PRCA will clearly suffice. The question is, of course, which one?

The typical process in the examination of multi-item scales is to sum the items and dichotomize the responses to obtain the scores of those who are high and those who are low on the particular scale. It is possible that the choice of usual scoring vs. recoded scoring would make no difference at all. Hence, the query was: would subjects be grouped differently if we summed and dichotomized the usual 5-4-3-2-1 scoring versus using the subjects' psychological scale values? As can be seen in Table 7, which compares the two versions for all subjects, the

---

Insert Table 7 about here.

---

two scoring methods do produce rather different answers. It is interesting to note that the self-esteem index has more than twice as many subjects (21%) differently classified between the two versions as the PRCA does (10%). The clear implications for the present study are that at least one person in ten would be erroneously classified if the "wrong" scoring method were chosen.

We chose to use the summed, dichotomized indices based on the recoded PRCA and self-esteem measures using the subjects' individual scale values for basically two reasons. First, as pointed out earlier, this scoring technique has less measurement error than the ordinary procedure. Second, we know the recoded measures are true interval values, so we may sum them with more confidence that the resulting composite indices are meaningful. Consequently, the analysis which follows concentrates on examining the subjects' psychological scale values on the PRCA and self-esteem.

It should be obvious from the discussion above that the dependent variables in this study are the 32 agreement scale response terms, the 21 evaluation scale response terms, and the 16 frequency scale response terms, for a total of 69 dependent variables. When one considers the nature of these 69 variables, the problem is clear albeit complex. If the dependent variables are highly intercorrelated, and in this case each of the three sets of adjectives are obviously highly related, a method must be found to account for the fact that we are examining multiple measures of an underlying common metric.

Analysis: One way to attack such a problem is via regression. Each of the 69 dependent variables is examined and the most predictive variable is selected to be included as the first term in an equation describing the linear predictive relationship between dependent and independent variables. Then, controlling for the effects of the variable(s) in the equation, the next best variable is selected and so forth. Since the explained variance is subtracted for each variable before the next variable is entered into the equation, only that portion of the variance remaining and not yet explained is taken into account for each

successive set. Hence, for practical purposes, multicolinearity poses few problems. The process continues until an arbitrary cut-off point (usually the .05 level of confidence) is reached. The general logic of regression is applicable to a more sophisticated technique allowing more than two categories for the exogenous variable as well as multiple endogenous variables: discriminant analysis. This technique is the statistic of choice here. Since the computer program packages for data analyses (Nie et al., 1975; Helwig and Council, 1979) routinely, and with good reason, exclude all subjects with any missing data from multivariate computations, this part is based on only 91 subjects.

The creation of a single variable on the personality dimension is straightforward since both communication apprehension and self-esteem are dichotomized. In all, we constructed four mutally exclusive groups. In group 1 are all the subjects having high communication apprehension and low self-esteem (N = 32); group 2 contains all those subjects with high communication apprehension and high self-esteem (N = 12); group 3 is composed of those low on communication apprehension and low on self-esteem (N = 13); and group 4 is comprised of those low in communication apprehension and high on self-esteem (N = 34).

The step-wise discriminant analysis using the minimum Wilks' lambda criterion with the four groups defined above as the single exogenous variable and the 69 scale response terms as the endogenous variables yielded a subset of 34 items which represented the significant variability in the endogenous variables (Table 8). The discriminant analysis results leads us to the

---

Insert Table 8 about here.

---

conclusion that there are significant differences in the way different personality types use the response categories of agreement, evaluation, and frequency. This knowledge, however, does not tell us how the groups differ from each other. To answer this further question, the canonical discriminant functions were examined.

Canonical discriminant functions can be used to study the spatial relationships among the four groups since they can be thought of as the axes of a geometric space displaying maximum differentiation among the groups. Because we have three functions, one less than the number of groups, the solution is essentially a geometric three-dimensional space presentation; such a solution is difficult to conceptualize and even more difficult to present. Were there only two dimensions, it would be simple to graph the solution since the subjects would be the points of intersection of their score values on each of the two functions.

With parsimonious hope we examined the computed eigenvalues, and their related percent of variance and canonical correlation to check the relative ability of each function to separate the groups. Also computed was the change in Wilks' lambda plus the associated chi-square value and significance level as

the information of each function is statistically removed in a step-wise manner from the overall discriminant space. As shown in Table 9, the first two

---

Insert Table 9 about here.

---

canonical discriminant functions together account for 82.95% of the variance. Since a small Wilks' lambda indicates more discriminating power, with all the data present Wilks' lambda is 0.067--a significant discrimination among the full set of items. After the second canonical function has been removed, the value of lambda is 0.547 and the significant variation is gone. With only the third function remaining in the analysis, the significance of the chi-square test is only 0.096. Essentially, then, we may safely deal with only the first two functions.

In Figure A one may visually examine the differentiation among the four

---

Insert Figure A about here.

---

groups that the first two canonical discriminant functions produce. As illustrated, groups 1 and 2 differ markedly from each other and from the other two groups in how they construct and use the scales. The fact that the first two functions do not clearly differentiate between groups 3 and 4 should not be taken to imply that these two groups are essentially the same in the psychological construction of their response scales. It must be recalled that groups 2 and 3 both have small samples (12 and 13 respectively) and that they constitute theoretical anomalies: high self-esteem apprehensives and low self-esteem confident communicators. It is clear from the data here that the first canonical discriminant function most clearly differentiates between groups 1 and 4.

Interpretation: This finally brings us to the point of trying to answer the question of how scales differ among different personality types. To deal with the question, correlations between each of the 69 response terms and the first two canonical discriminant functions were calculated for all of the four groups. The correlations are presented in Tables 1, 2, and 3. In general these correlations may be used as an index of stability and instability. Since the first two canonical functions indicate maximum differentiation among the groups, those items which most highly correlate with the first two functions (either positively or negatively; sign does not matter here) are the items with maximum variability among the groups. To find the stable items, one simply searches for items having the smallest absolute value correlations with the two functions.

In examining Tables 1, 2, and 3, it should first be noticed that the standard deviations for all the items are generally higher for group 1 than for group 4.

These results would seem to imply that the low self-esteem apprehensives are in less agreement about the meaning of the terms than the high self-esteem confident communicators. In fact, this observation may be a hallmark of self-esteem and communication anxiety. A person adrift in a sea of fluid values might be, quite obviously, unsure of which way to turn or what standard to trust. Tentativeness and caution would be sensible courses of action in a universe of chaotic values. Viewed in this light, what better backdrop could we offer for a low self-esteem communication apprehensive? On the other hand, a social world of stability and predictability would probably be equally descriptive of the psychological environment of a high self-esteem confident communicator.

Having, in general, described the differences in orientation between low self-esteem communication anxious persons (group 1) and high self-esteem communication confident persons (group 4) perhaps we can shed more light on the mechanism by which these individuals reach such an orientation. Focusing, for example, only on the negative items of "Rarely," "Disagree," and "Very much disagree," it may be seen that all the standard deviations for these terms are smaller for group 4 than for group 1; however, it is in the pattern of the means where the real clues seem to lie. For "Rarely," group 4 has assigned a mean value of 9.18 while group 1 gave it 18.16; for "Disagree," group 4 gave it a mean value of 13.15 while group 1 gave it 27.22; and for "Very much disagree," group 4 gave it a value of 5.12 while group 1 gave it 14.25. These results suggest that low self-esteem apprehensives tend to pull their scales up, to move the negative terms closer to the positive end of the scale than is the case among high self-esteem confident communicators. In a similar pattern among the positive ends of the scales, the group 4 mean values are consistently higher than the group 1 mean values. That is, the positive items are pulled down toward the negative end for the low self-esteem apprehensives more so than for the high self-esteem communicators. In essence, the low self-esteem communication anxious person may exhibit such traits because of the psychological limits of his or her own perceptual system. Many of us can sympathize with the torture a person must feel when physical walls slowly close around them. Surely, the world is no less terrifying when the conceptual walls similarly inch inward.

In the final phase, a good deal of information may also be learned about scales and scaling per se by careful inspection of Tables 1, 2, and 3. For example, if the scales are divided roughly into high, middle and low ranges, the correlations seem to indicate that the center regions of the scales are more stable than the ends. This is the easiest to see with the evaluation scale (Table 3). The top seven response terms have five correlations with an absolute value above 0.10, the bottom seven response terms have two correlations with an absolute value above 0.10, while the middle seven have none.

Another implication to be drawn is to note which items are the most variable from all the scales. A search of Tables 1, 2, and 3 reveals that four items have correlations with an absolute value greater than 0.20. They are: "Rarely" from the frequency scale (Table 3); "Disagree" from the agreement scale (Table 1); "Terrible" from the evaluation scale (Table 2); and finally, "Very much disagree" from the agreement scale (Table 1). In all cases, these are negative terms. This observation leads us to the generalization that the negative ends of the scales may be more unstable than the positive ends.

Implications

Retreating to examine the study as a whole, a number of implications are apparent. What has been identified for many years as "measurement error" in scaling circles may now be attributed, at least in part, to a failure to allow subjects the freedom to define their own perceptual universe. It is probably not the case that the social sciences must start over but we may now be a bit more precise than we have been in the past.

Second, as if to compound the problem of scaling, it is equally clear that an additional source of variation lurks in the personality orientations among subjects. We would not be so presumptuous as to suggest that self-esteem and communication anxiety are the two most critical variables to be taken into account, but they are two which apparently make a difference.

In the final analysis, it is our hope that more diligent researchers will find further refinements in this general line of inquiry but that we have begun to illuminate a few dark corners with this small study.

Footnotes

[1]"Barely passing" was not measured in the present study. The value shown is for the descriptor "Passable."

References

Acito, F., & Anderson, R. D. A monte carlo comparison of factor analytic methods. Journal of Marketing Research, 1980, 17, 228-236.

Allport, G. W., & Pettigrew, T. F. Cultural influence on the perception of movement: the trapezoidal illusion among Zulus. Journal of Abnormal and Social Psychology, 1957, 55, 104-113.

Beatty, M. J., Behnke, R. R., & McCallum, K. Situational determinants of communication apprehension. Communication Monographs, 1978, 45, 187-191.

Gordon, C., & Gergen, K. The Self in Social Interaction. New York: Wiley, 1968.

Hastorf, A. H., & Cantril, H. They saw a game: a case study. Journal of Abnormal and Social Psychology, 1954, 49, 129-134.

Helwig, J. T., & Council, P. S. (eds.), SAS Users Guide, 1979. Raleigh, North Carolina: Statistical Analysis System Institute, 1979.

Hensley, W. E. Differences between males and females on the Rosenberg scale of self-esteem. Psychological Reports, 1977, 41, 829-830.

Hensley, W. E., & Roberts, M. K. Dimensions of Rosenberg's self-esteem scale. Psychological Reports, 1976, 38, 583-584.

Ittelson, W. H., & Kilpatrick, F. P. Experiments in perception. Scientific American, 1951, 185, No. 2, 50-55.

McCroskey, J. C. Measures of communication-bound anxiety. Communication Monographs, 1970, 37, 269-277.

McCroskey, J. C. Oral communication apprehension: A summary of recent theory and research. Human Communication Research, 1977, 4, 78-96.

McCroskey, J. C. Validity of the PRCA as an index of oral communication apprehension. Communication Monographs, 1978, 45, 192-203.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. Statistical Package for the Social Sciences. New York: McGraw-Hill, 1975.

Parks, M. R. A test of the cross-situational consistency of communication apprehension. Communication Monographs, 1980, 47, 220-232.

Robinson, J. P., & Shaver, P. R. Measures of Social Psychological Attitudes. Ann Arbor, Mich.: Survey Research Center, 1969.

Rosenberg, M. Society and the Adolescent Self-Image. Princeton, N. J.: Princeton University Press, 1965.

Spector, P. E. Choosing response categories for summated rating scales. Journal of Applied Psychology, 1976, 61, 374-375.

Table 1

Agreement Scale

Response Categories	Overall MEAN	Overall S.D.	Group 1 MEAN	Group 1 S.D.	Group 2 MEAN	Group 2 S.D.	Group 3 MEAN	Group 3 S.D.	Group 4 MEAN	Group 4 S.D.	CORRELATIONS Fun 1	CORRELATIONS Fun 2
Strongly agree	90.95	14.28	86.31	20.70	94.33	5.84	87.77	17.92	94.32	6.15	0.134	0.013
Very much agree	88.71	16.15	84.00	22.09	90.17	6.13	87.23	22.03	94.29	5.46	0.116	-0.085
Agree on the whole	84.20	15.39	81.53	16.00	87.67	10.00	72.54	23.94	87.56	11.44	0.067	0.048
Agree	83.95	15.69	82.34	19.10	88.58	9.90	79.23	18.28	87.35	10.33	0.088	-0.007
Pretty much agree	75.94	13.21	73.75	18.10	71.50	10.60	73.46	12.69	78.97	7.43	0.054	-0.161
Generally agree	71.09	14.16	69.72	18.28	68.50	9.65	68.69	14.80	74.03	10.93	0.053	-0.123
Inclined to agree	66.51	12.46	65.56	14.89	64.08	10.47	63.92	9.76	70.12	11.56	-0.026	-0.124
Agree in part	65.69	11.68	64.97	13.16	68.17	6.91	66.08	11.49	64.50	11.28	0.117	0.166
Tend to agree	65.13	9.66	64.47	11.13	60.25	8.85	66.31	7.63	65.65	8.19	-0.001	-0.186
Moderately agree	64.93	10.94	65.06	14.43	64.75	9.74	61.92	9.05	66.38	7.45	-0.010	-0.024
Halfway agree	59.01	12.30	57.69	13.17	62.42	11.77	55.77	12.94	59.00	11.15	0.069	0.053
Mildly agree	58.83	10.60	57.75	11.91	56.33	9.81	57.69	9.48	59.53	10.58	0.018	-0.017
Somewhat agree	57.91	10.89	56.31	12.33	59.58	8.77	58.15	12.19	58.76	11.47	0.062	0.020
Slightly agree	53.24	12.70	50.25	16.13	49.67	12.02	51.46	12.86	56.03	9.19	0.077	-0.144
Agree a little	52.90	14.36	51.72	17.60	50.17	18.10	52.15	14.01	54.56	12.18	0.026	-0.084
Indifferent	50.46	8.86	51.41	12.41	50.83	0.39	47.15	12.44	50.06	5.14	-0.051	0.046
Undecided	49.44	11.79	50.22	15.80	46.58	14.37	47.15	13.57	51.12	0.69	-0.020	-0.102
Slightly disagree	45.34	10.77	47.28	13.08	45.08	11.97	43.62	12.76	43.62	7.77	-0.078	0.049
Disagree a little	44.79	11.74	45.84	14.52	48.08	17.11	44.08	12.02	42.41	7.46	-0.041	0.141
Mildly disagree	42.23	9.17	42.22	12.83	42.00	8.52	41.77	7.66	41.94	5.47	0.085	-0.109
Somewhat disagree	41.26	11.52	45.84	11.54	36.08	8.15	36.77	15.25	39.50	11.00	-0.187	-0.067
Halfway disagree	38.67	11.63	40.72	13.57	38.25	11.10	37.00	10.75	37.44	10.02	-0.127	0.050
Tend to disagree	38.65	14.04	41.53	14.75	35.75	10.20	42.69	18.54	35.18	12.90	-0.100	-0.007
Disagree in part	38.39	13.39	41.38	15.15	37.33	11.31	42.08	15.60	34.24	11.62	-0.107	0.051
Inclined to disagree	35.85	12.63	38.09	14.81	34.00	12.18	38.38	9.87	32.74	11.24	-0.014	-0.025
Moderately disagree	35.43	11.31	35.69	15.36	34.25	11.53	36.54	7.16	35.24	8.45	-0.012	-0.032
Generally disagree	26.14	14.22	29.41	20.80	24.25	10.48	24.46	9.44	21.94	8.64	-0.128	0.051
Pretty much disagree	21.84	13.78	21.19	14.66	21.83	11.82	20.15	6.41	21.35	14.49	-0.002	0.068
Disagree	18.10	17.33	27.22	24.19	9.50	8.19	16.77	9.68	13.15	9.62	-0.250	-0.063
Disagree on the whole	17.37	19.15	24.78	26.00	11.33	9.89	22.69	23.57	12.71	10.71	-0.158	-0.037
Strongly disagree	9.49	15.09	15.09	22.22	5.75	3.79	13.54	19.93	5.59	4.47	-0.172	0.026
Very much disagree	8.37	11.48	14.25	18.44	6.25	4.71	7.00	4.30	5.12	4.23	-0.210	0.038

Note. Each item of the Overall column is based on from 101 to 103 subjects. The number of subjects in Group 1 is 32; in Group 2 it is 12; in Group 3 it is 13; and in Group 4 it is 34. Discrepencies are due to missing data.

Table 2

Evaluation Scale

Response Categories	Overall MEAN	Overall S.D.	Group 1 MEAN	Group 1 S.D.	Group 2 MEAN	Group 2 S.D.	Group 3 MEAN	Group 3 S.D.	Group 4 MEAN	Group 4 S.D.	CORRELATIONS Fun 1	CORRELATIONS Fun 2
Excellent	94.91	9.90	91.16	16.29	96.58	5.21	95.46	4.14	96.79	3.67	0.133	0.022
Superior	94.39	11.41	90.19	17.64	97.00	4.73	95.00	8.12	96.38	6.44	0.136	0.012
Outstanding	93.25	11.24	89.50	18.06	97.50	2.81	95.23	4.76	93.94	6.48	0.134	0.081
Very good	85.46	10.87	81.09	15.48	86.33	8.72	88.85	5.24	87.15	7.97	0.157	-0.043
Above average	75.58	10.14	75.25	11.15	79.00	12.68	76.46	8.03	74.35	9.22	-0.086	0.147
Good	74.80	12.41	73.06	14.05	75.50	10.84	76.77	13.05	75.50	10.23	0.062	-0.010
Decent	63.10	14.52	62.00	18.03	64.67	11.66	59.62	13.35	64.44	12.34	0.033	0.019
Satisfactory	60.73	13.02	61.09	12.08	59.75	15.74	54.31	15.75	60.62	11.18	-0.038	0.007
Acceptable	60.46	13.06	58.88	16.19	59.33	11.15	60.23	11.34	61.44	12.51	-0.068	-0.024
Average	57.06	11.93	55.81	15.60	57.25	11.31	54.46	8.01	57.62	9.67	0.028	0.002
Fair	56.44	14.86	55.28	16.67	57.00	16.38	60.08	12.43	56.44	11.73	0.092	0.002
Passable	53.24	13.83	50.78	16.12	56.42	12.87	55.92	9.53	53.56	12.70	0.080	0.051
Mediocre	52.48	13.68	49.75	16.36	49.75	17.41	51.15	13.27	53.94	9.26	0.056	-0.090
Needs improvement	49.56	16.84	49.88	19.58	51.08	20.14	46.69	10.59	49.47	15.00	-0.008	0.039
Unsatisfactory	36.09	16.83	36.66	16.18	36.75	21.93	36.08	16.55	34.35	17.64	-0.037	-0.029
Poor	32.37	16.45	32.56	15.78	36.75	21.76	29.54	16.99	34.26	15.97	0.033	-0.009
Inferior	31.28	19.63	32.59	23.06	34.83	17.55	28.15	17.06	29.62	19.17	-0.032	-0.039
Bad	28.87	16.00	30.00	17.35	28.75	20.17	26.00	17.91	27.62	13.05	-0.003	0.127
Failing	27.10	21.54	25.25	25.29	29.58	22.12	31.62	21.19	25.85	17.86	0.032	0.038
Very poor	23.06	16.91	24.47	17.31	27.33	23.48	19.46	16.63	22.97	15.10	-0.017	0.091
Terrible	16.67	15.51	15.31	15.28	27.17	22.18	17.31	16.32	15.71	13.01	-0.071	0.224

Note. Each item of the Overall column is based on from 101 to 103 subjects. The number of subjects in Group 1 is 32; in Group 2 it is 12; in Group 3 it is 13; and in Group 4 it is 34. Discrepencies are due to missing data.

Table 3

Frequency Scale

Response Categories	Overall MEAN	Overall S.D.	Group 1 MEAN	Group 1 S.D.	Group 2 MEAN	Group 2 S.D.	Group 3 MEAN	Group 3 S.D.	Group 4 MEAN	Group 4 S.D.	CORRELATIONS Fun 1	CORRELATIONS Fun 2
Always	91.30	21.75	83.25	29.92	98.50	2.81	94.46	8.86	98.50	2.81	0.190	-0.022
Most of the time	79.83	13.10	78.47	12.80	82.67	11.10	80.08	11.34	80.62	12.92	0.100	0.035
Repeatedly	76.28	17.87	73.44	16.91	84.75	9.94	75.92	18.29	80.62	14.01	0.064	-0.016
Frequently	74.42	14.81	71.16	17.06	72.67	16.28	74.69	9.34	77.85	8.08	-0.013	-0.098
Often	69.93	11.21	67.13	10.51	73.25	7.71	68.00	13.15	70.32	11.66	0.090	0.088
Pretty often	69.81	16.47	65.09	20.28	76.00	8.51	69.62	18.44	72.06	13.54	0.129	0.073
Sometimes	47.58	15.79	46.56	16.55	45.92	9.44	48.31	14.38	46.35	17.93	-0.084	0.031
Now and then	41.00	16.42	43.88	15.45	36.67	15.69	37.77	14.08	40.62	16.92	-0.050	0.035
Occasionally	40.53	15.98	41.66	14.91	38.00	15.27	36.69	16.60	40.56	15.72	-0.119	-0.012
Once in a while	31.27	16.74	33.28	17.18	27.83	17.70	31.08	18.95	29.68	13.72	-0.082	-0.014
Infrequently	25.21	10.88	27.19	11.67	18.33	7.44	22.69	7.23	25.29	9.32	-0.052	-0.084
Seldom	21.51	12.54	25.00	12.08	18.50	13.14	18.62	10.00	21.06	14.44	-0.109	-0.045
Pretty seldom	21.32	10.76	25.13	11.65	20.58	9.19	19.54	9.73	17.79	8.79	-0.180	0.092
Once or twice	17.80	15.74	20.16	17.23	9.42	4.80	21.15	24.62	15.76	12.13	-0.167	-0.144
Rarely	13.68	12.52	18.16	16.26	14.92	15.62	11.00	6.43	9.18	5.73	-0.259	0.105
Never	7.07	15.98	11.63	21.30	2.08	1.93	4.69	4.68	5.62	16.97	-0.120	-0.056

Note. Each item of the Overall column is based on from 101 to 103 subjects. The number of subjects in Group 1 is 32; in Group 2 it is 12; in Group 3 it is 13; and in Group 4 it is 34. Discrepencies are due to missing data.

Table 4

Varimax Rotated Factor Matrices of the Usual vs. Recoded Rosenberg Self-esteem Measure

	Usual Version		Recoded Version	
Item	Factor 1	Factor 2	Factor 1	Factor 2
1	-0.36451	0.59311	-0.13758	0.67591
2	0.60454	-0.42700	0.62581	-0.30056
3	-0.67123	0.36806	-0.38044	0.61351
4	-0.18211	0.61721	-0.09539	0.60052
5	-0.20015	0.64696	-0.00647	0.61390
6	0.76172	-0.17061	0.75823	0.01400
7	0.34192	-0.70328	0.62796	-0.28787
8	-0.70432	0.44005	-0.40905	0.69847
9	0.66828	-0.25609	0.80414	-0.12853
10	0.27372	-0.43411	0.57625	-0.14031

Table 5

Varimax Rotated Factor Matrix of PRCA
(Usual Version)

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
1	-0.16931	0.09942	0.58904	0.00468	0.18888
2	0.53703	-0.07755	-0.24783	0.21326	0.06639
3	-0.19669	0.25775	0.45993	-0.25441	0.15093
4	0.38055	-0.10616	-0.20437	0.70752	-0.14530
5	-0.03582	0.25787	0.45317	-0.16569	0.37899
6	0.17049	-0.21338	-0.15133	0.74696	-0.27175
7	0.61393	-0.04473	0.04418	0.08135	-0.28450
8	-0.06435	0.46310	0.45274	-0.02619	0.04408
9	-0.24837	0.44978	0.48967	-0.22971	0.26232
10	-0.42491	0.35587	0.16915	-0.18890	0.28741
11	0.18746	-0.15335	-0.51530	0.23927	-0.11383
12	-0.41745	0.37667	-0.11711	-0.18472	0.27703
13	-0.55676	0.20794	0.18153	-0.17039	0.24179
14	0.47943	-0.16541	-0.49327	0.17167	0.04437
15	-0.53028	0.47078	0.22640	-0.23097	0.04141
16	-0.40115	0.57621	0.25410	-0.10065	0.27916
17	0.30864	-0.42658	-0.11690	0.44775	0.00262
18	0.53480	-0.34517	-0.28795	0.02741	-0.16096
19	-0.11490	0.01811	0.31965	-0.02577	0.52625
20	-0.26243	0.42947	0.17719	-0.18072	0.00614
21	0.38253	-0.50421	-0.27579	0.22199	-0.21696
22	-0.19748	0.17179	0.07778	-0.21575	0.05302
23	0.56624	-0.20794	-0.26594	0.22547	-0.06231
24	-0.04231	0.46468	0.27345	-0.12533	0.37224
25	0.40686	-0.30018	-0.22064	0.17065	-0.10378

Table 6

Varimax Rotated Factor Matrix of PRCA
(Recoded Version)

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
1	0.33044	0.20787	-0.00943	0.04817	-0.09895	0.65392	0.10362
2	0.16470	-0.11864	0.13909	0.07236	0.58270	-0.09823	0.00198
3	0.28932	0.53122	-0.10192	-0.04254	-0.12120	0.17000	0.11573
4	-0.04856	-0.15375	0.88048	0.12532	0.29017	-0.05342	0.00172
5	0.70020	0.26761	-0.06664	-0.00634	-0.06717	0.05840	0.05985
6	-0.17931	-0.11832	0.79291	0.21592	0.09847	0.02324	-0.09108
7	0.08096	-0.08316	0.43699	0.32947	0.13585	-0.15855	-0.27003
8	0.47048	0.06805	-0.01021	-0.16831	0.12622	0.47292	0.19577
9	0.70349	0.31250	-0.14897	-0.16156	-0.26012	0.11294	0.02914
10	0.15404	0.74445	-0.09503	-0.06974	-0.11699	0.07912	0.11430
11	-0.24162	-0.02111	0.21179	0.06580	0.58215	0.01149	0.02946
12	0.22628	0.46463	0.00903	-0.19818	-0.09749	-0.17697	0.22937
13	0.28360	0.49141	-0.21320	-0.29182	-0.02134	0.13263	0.05283
14	-0.06449	-0.09687	0.30936	0.23972	0.36730	-0.42894	0.07223
15	0.40876	0.51537	-0.09443	-0.08929	-0.03915	0.10569	0.01938
16	0.43222	0.54000	-0.12312	-0.18514	-0.08888	0.19383	0.01509
17	-0.13274	-0.11051	0.32838	0.58691	0.12948	0.23651	0.00058
18	-0.20122	-0.06810	0.16484	0.57095	0.35757	-0.17289	-0.00788
19	0.36246	0.12332	0.06906	0.09506	-0.09395	0.22487	0.34342
20	0.39556	0.26203	-0.03258	-0.28169	0.12484	0.09648	0.12044
21	-0.08471	-0.17847	0.06223	0.55420	0.38046	-0.05507	-0.24158
22	0.18464	0.20655	-0.16373	-0.07411	0.00283	0.06007	0.82494
23	0.04331	-0.07988	0.03227	0.25911	0.57890	0.02072	-0.08315
24	0.54887	0.20351	-0.05154	-0.09002	0.15531	0.11066	0.10329
25	-0.03072	-0.30120	0.19765	0.52951	0.12496	-0.07804	0.07668

Table 7

Comparison of How Summed, Dichotomized Self-esteem and PRCA Measures Using Usual vs. Recoded Response Values Catagorize the Subjects

Usual Self-esteem	Recoded Self-esteem		Usual PRCA	Recoded PRCA	
	high self-esteem	low self-esteem		high apprehension	low apprehension
high self-esteem	40	10	high apprehension	46	5
low self-esteem	12	41	low apprehension	5	47

Table 8

Multivariate Analysis of Variance of Self-esteem and Apprehension with Subset of 34 Most Variable Items

Exogenous Variables	Hotelling-Lawley Trace	Pillai's Trace	Approximate F (D.F. = 34, 61)	Significance
Self-esteem	1.496	0.599	2.68	0.000
Apprehension	1.122	0.529	2.01	0.009
Interaction	0.989	0.497	1.78	0.025

Table 9

Tests of Relative Differentiability of
Canonical Discriminant Functions

Function	Eigenvalue	Percent of Variance	Canonical Correlation	After Function	Wilks' Lambda	Chi-Squared	D. F.	SIGNIFICANCE
1	2.937	60.51	0.864	0	0.067	192.420	102	0.000
2	1.089	22.44	0.722	1	0.262	95.124	66	0.011
3	0.828	17.05	0.673	2	0.547	42.811	32	0.096

FIGURE A

ALL-GROUPS SCATTERPLOT - *INDICATES A GROUP CENTROID

CANONICAL DISCRIMINANT FUNCTION 1